Reinventing Security for Trust in Cyberspace*

Fred B. Schneider†

Department of Computer Science
Cornell University
Ithaca, New York 14853

1 Introduction

A National Research Council CSTB study concludes, in its final report *Trust in Cyberspace* [1], that we lack the science and technology base for building trustworthy networked information systems. A *trustworthy system* by definition does what its designers intend, which invariably means it must tolerate environmental disruption, operational errors, the inevitable design errors, and hostile attacks. This note (much of whose content is drawn from [1]) summarizes a banquet speech about one of those dimensions: security.

Why focus on security, when no significant critical infrastructure outages have been attributed to security breaches? The reason is simple. Attacks are increasing at the same exponential rate as the internet. And this trend should not be surprising. There is growing incentive for attackers as infrastructures come to depend on networked computers and as these are interconnected. Attackers see "one-stop" shopping (from a workstation) and an ever-expanding payoff for success.

[*] Summary of a banquet address given as part of the conference *Beyond Moore's Law, Opportunities and Threats from Future, Ubiquitous, High-Performance Computing*, Center for Global Security Research, Lawrence Livermore National Laboratory, December 8–10, 1999.

[†] Supported in part by ARPA/RADC grant F30602-96-1-0317, AFOSR grant F49620-94-1-0198, Defense Advanced Research Projects Agency (DARPA) and Air Force Research Laboratory, Air Force Material Command, USAF, under agreement number F30602-99-1-0533, National Science Foundation Grant 9703470, and a grant from Intel Corporation. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of these organizations or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright annotation thereon.

2 Nature of The Technical Problems

Computer security is not a new concern. But two things are new about today's problem: the way systems are now constructed and the types of security now required.

System Construction. Systems today are built from COTS (common off the shelf) components and subsystems, because this reduces project costs and risks. But the system builder who includes COTS software in a networked information system cannot know exactly what is in that system nor how it works. The system builder also becomes dependent on a third party for change and fixing (certain) bugs. Finally, the limited access to internal interfaces of COTS software implies that so-called non-functional properties, like security and reliability which are ultimately defined in terms of internal interfaces of components, are impossible to test.

That COTS development tends to be driven by market-entry timing and customer requests for new features means that correctness and assurance are not paramount concerns for COTS developers. Thus, the system builder who employs COTS components works with building blocks that are likely to have residual errors and therefore are likely to have vulnerabilities. COTS components also are increasingly built to be extensible. This allows the functionality of these components to be enhanced after they are fielded, which helps a component's producer preserve market share (since a customer can enhance the existing component more easily than switching to another). It is hard enough to build a networked information system that works as delivered—now the system must continue to work as it evolves in unimagined ways.

Networked information systems are almost never conceived and built from the ground up. Instead, these systems grow by accretion and agglomeration. Substantial legacy contact is thus the norm, where typically nobody has a good understanding of how or why all the legacy components work. In addition, networked information systems have scope that transcends national and corporate boundaries. This geographic and political scope means that no single authority is empowered to control the growth or evolution of the system. Subsystems cannot completely trust each other but must cooperate despite this mutual distrust, requiring new approaches to system organization.

Security Needs. Historically, computer and communications security has been dominated by concern for enforcing secrecy properties. For systems intended to control critical infrastructures, availability and integrity assume increased importance. The “Principle of Least Privilege” is crucial for protecting against errant system extensions and foreign code, but applying this principle requires fine-grained access control, something that virtually no operating systems today support, and requires access rights determined as much by past actions as by the current system state. A policy to control a remotely executing software-upgrade service, for example, might stipulate “No sends are allowed after reading from any but a few configuration files” — and this cannot be enforced by the read/write/execute permission bits supported in today’s operating systems.

The scale and geographic scope of networked information systems also causes new problems in implementing security. Cryptography can help in securing communications channels and in maintaining the integrity of network routings. But there is no experience with key-management infrastructures for systems the size of the Internet. Revocation and recovery from compromised keys, for example, are relatively easy at small-scales but hard to manage when the number of principals gets very large. Name-space management is similarly affected by scale. Not just the number of principals but also the absence of a single trusted authority to control a network-wide resource (like keys or names) leads to new and unsolved problems.

3 Nature of the Non-Technical Problems

Networked information systems are complex, and it is the nature of complex systems that they defy understanding. Since a system that is not understood is likely to contain vulnerabilities, absolute security is out of reach for networked information systems. With knowledge of threats, however, we can direct investment to increase assurance or to add defensive layers tailored for those anticipated threats. But doing so requires a paradigm shift and a resetting of expectations—rejecting the dogma of absolute security and embracing risk-management.

Society as a whole and people as individuals will sanction investments in trustworthiness only to the extent that such investments can be seen to reduce risks.[1] Information about the probabilities and costs of breaches is thus required before significant investments for risk reduction should be expected. Unfortunately, it is difficult to identify, much less quantify, risks and costs associated with security breaches. One must somehow ascertain

[1] Increased trustworthiness does, in some cases, open new markets or create new business opportunities. Banking and e-commerce are examples.

whether the system is an attractive enough target, which is a function of the potential payoff and the potential effort in waging a successful attack. And people cannot even build good intuitions about the frequency of successful attacks when the institutions most likely to be attacked (banks and the military) have considerable disincentive to report compromises and remain silent about such matters.

Although direct costs of security breaches might be understood, indirect costs—often the more consequential—are notoriously hard to quantify or even identify. An attack that causes a telephone outage obviously deprives a phone company of revenue, but the attack also might isolate burglar alarms from police stations and citizens from police and medical protection—potentially significant costs that the telephone provider does not incur. Similarly, the implications of broadcasting somebody's private medical records could range from embarrassment to altering the outcome of a national election. What is the cost of such a breach?

Were there a real need for security, one might expect there to be a thriving market for security goods and services. The absence of such a market might then be taken to mean there is no need for security. This argument is flawed, however. A market works only if buyers and sellers have information about costs and benefits of the commodities being exchanged. Such information about security is largely unavailable. For example, assurance technology often makes a system easier to understand and debug—what portion of such a technology's cost should be assigned to security? Other trustworthiness-enhancers are likely to delay time-to-market for a product. How can that cost be measured?

4 The Future

This is not our first love affair with technology. Looking back over the last 50 years, we see:

- In 1957, the first American nuclear power plant went into operation. Hopes of clean, cheap energy were subsequently dashed with incidents at Three-Mile Island and Chernobyl.
- In 1962, Rachel Carlson's *Silent Spring* started an international movement to restore the quality of our environment. Slow and steady progress ensured, with noticeable dividends now apparent.

Today, we are in the midst of a love affair with computers and networking. Let to continue, our society's infrastructures will become dependent on networked information systems. Disrupt cyberspace and you disrupt society, so a series of computer system compromises could serve as a Three-mile Island or Chernobyl wake-up call. On the other hand, a wide understanding of the risks might lead to investments that prevent such a disaster. While the proximate causes of the predicted disasters are technological, note that any solution will be driven not by technologists but rather by a mandate from society.

Were there the will, would there be a way? Today, we lack the technology to build networked information systems that are trustworthy enough to control critical infrastructures. New research is required. But once that reseach has been done, Moore's Law becomes important. Every 5 years or so, significant additional computational capacity becomes available to software system designers. In the past, this capactiy was devoted to providing graphical user interfaces (which expanded the potential market for computers by enlarging the user base) and programmer-productivity enhancing tools (which allowed applications to be written for that market). How future increases in computational power are spent has not been decided. Trustworthiness anyone?

Acknowledgments

Working with NRC CSTB study on Information Systems Trustworthiness was crucial in developing the ideas discussed in this note.

References

[1] Schneider, F.B. (ed) *Trust in Cyberspace*. National Academy Press, Washington, D.C. 1999.